RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: _09/855,340B_

Source: _IFW116_

Date Processed by STIC: _2-3-05_

ENTERED

RAW SEQUENCE LISTING   DATE: 02/03/2005
PATENT APPLICATION: **US/09/855,340B**   TIME: 09:26:38

Input Set : **A:\PTO.ST.txt**
Output Set: **N:\CRF4\02032005\I855340B.raw**

```
 3 <110> APPLICANT: Hosted, Jr., Thomas J.
 4        Horan, Ann C.
 6 <120> TITLE OF INVENTION: Isolation of Micromonospora carbonacea var africana
 7        pMLP1 integrase and use of integrating function for
 8        site-specific integration into Micromonospora
 9        halophitica and Micromonospora carbonacea chromosome
11 <130> FILE REFERENCE: IN01164K
13 <140> CURRENT APPLICATION NUMBER: 09/855,340B
14 <141> CURRENT FILING DATE: 2001-05-15
16 <150> PRIOR APPLICATION NUMBER: 60/204,670
17 <151> PRIOR FILING DATE: 2000-05-17
19 <160> NUMBER OF SEQ ID NOS: 16
21 <170> SOFTWARE: PatentIn Ver. 2.1
23 <210> SEQ ID NO: 1
24 <211> LENGTH: 1179
25 <212> TYPE: DNA
26 <213> ORGANISM: Micromonospora carbonacea
28 <400> SEQUENCE: 1
29 gtgtggatcg agaagaacgg gcccgtctac cgcattcggg acctcgttcg cggtaaaaag 60
30 gtcaccattc agaccggtta tccgacgaag accagcgcca agaatgcgat ggtgcagttc 120
31 cgtgcggagc agttgcaggg caacgcgctc atgccgcgcg gcggtcagat taccctcgcc 180
32 gatttcgtgg gggagtggtg gccgagctac gaaaagacgc tgaaaccgac cgccgtgaac 240
33 tcggagggca accggatccg caaccacctc ctgcccatac tcggccatct caccccttgac 300
34 gagctggacg ggcaggtcac ccagcagtgg gtcaacgacc tggaggccgg cgtcggcccg 360
35 tggccggagt ccacgcgggg tcgtcggaag ccgctggcag cgaagacgat cagcaactgc 420
36 cacggcctgc tgcacacgat ctgcggcgcg gcgatcgcgg cgaaacggat caggctcaac 480
37 ccgtgctctt cgacgatgct gccccggcgc gagccgaaag agatgaagtt cctgagcgac 540
38 ccggagatcg gtcggcttat cacggcgctt ccgccgcact ggcgaccgct cgtcatgctg 600
39 ctggtggcga ccggtctgag gtggggtgag gcgatcggcc tgcgcgccgg ccgggtcgac 660
40 ctgctcgccg cgcggccccg gctgaccgtc gtcgagcagc tccaggagct ggccagcacg 720
41 ggagagctcg tcttccagtc gccgaagacc gcgaagggcc ggcgcacggt cagtttcacc 780
42 acgaaagtcg ctctactgct tacgccactc atcgccggaa agaaaagtga cgaggtcgtg 840
43 ttcaccgcgc cgaaaggcgg gatggtaagg acgcgcaatt ccggcggat ctgggtcaag 900
44 gcgtgcgagg aagccgggct tccgggctta cgcattcacg atctgcggca cactcacgcg 960
45 gcgatcctga tttctgccgg gcgtccgctg tcggcgatct cccgccgcct cggtcactcg 1020
46 tcgatcgcgg tcacggatct gctgtacggg cacctgcgtg aggaggtcga cgagggatc 1080
47 ctcgcggcga tcgaggaggc gatggccggc gtccgggctg aggacctgga ggcggaactc 1140
48 gacgaggagc tgacggacgt gttggccgac gcagcatga 1179
51 <210> SEQ ID NO: 2
52 <211> LENGTH: 426
53 <212> TYPE: DNA
54 <213> ORGANISM: Micromonospora carbonacea
56 <400> SEQUENCE: 2
```

RAW SEQUENCE LISTING   DATE: 02/03/2005
PATENT APPLICATION: **US/09/855,340B**   TIME: 09:26:38

Input Set : **A:\PTO.ST.txt**
Output Set: **N:\CRF4\02032005\I855340B.raw**

```
57 atgcgcaaca caccgggggct ggggcgcggc acatgggccg catacgtcct caccgcccgc 60
58 gagcgcgccg gactgaccaa gagcgagttg gccaggcgca tccagaagga ccgggccacc 120
59 gtcggccggt gggaggacgg caagaaccgg cccgacgacg cggacctcgt tgcccgcgtc 180
60 gcccaggtgc tcggcctcga cctcgacgaa gccctcgccg ccgcaggtct gcgcccgggc 240
61 gtcacccgc cagcgacccc aaccatggac ctggacgagg aaatcgagct ggtccgcacc 300
62 gacccaagc tggacgagga catgaagcgg cgcatcatcg ccctaatcct ggagcgccgt 360
63 gagcgcgaca aggcggcggc gatcgaggaa accaagcggc tcatcgacct gttccgccgg 420
64 agctga                                                          426
67 <210> SEQ ID NO: 3
68 <211> LENGTH: 34
69 <212> TYPE: DNA
70 <213> ORGANISM: Micromonospora carbonacea
72 <400> SEQUENCE: 3
73 ccccggtacg ggttcaattc ccatcagtca cccg                           34
76 <210> SEQ ID NO: 4
77 <211> LENGTH: 241
78 <212> TYPE: DNA
79 <213> ORGANISM: Micromonospora carbonacea
81 <400> SEQUENCE: 4
82 tattagtccg cacgccgccc ggccccgccg gagcggagcg catggtggct gtagctcagt 60
83 tggcagagca ccgggttgtg gtcccggttg tcgtgggttc aattcccatc agtcacccgt 120
84 acacgaaggc cccctccact cggaggggggc cttcggcgtt cctgagggtt cgcggtcagg 180
85 cggtcggctc ggcgctgggg gactcggccc cgtcggcggg agtggcctcg gcgtccgggg 240
86 a                                                               241
89 <210> SEQ ID NO: 5
90 <211> LENGTH: 243
91 <212> TYPE: DNA
92 <213> ORGANISM: Micromonospora carbonacea
94 <400> SEQUENCE: 5
95 tggcggggggt gtggctatta ttagtccgca cgccgcccgg ccccgccgga gcggagcgca 60
96 tggtggctgt agctcagttg gcagagcacc gggttgtggt cccggttgtc gtgggttcaa 120
97 ttcccatcag tcacccggca agtggatcta ctccacagca gatcaggccc cctccgaaga 180
98 gggggcctga tgcgtcatag gggacaggta ggggaactca accccggct ccttgctcgc 240
99 gtc                                                             243
102 <210> SEQ ID NO: 6
103 <211> LENGTH: 247
104 <212> TYPE: DNA
105 <213> ORGANISM: Micromonospora carbonacea
107 <400> SEQUENCE: 6
108 tagggggaatc cactccggag acgcccggag caatccggag catgacggag caaccagcag 60
109 gtcaggtggc ctgttgaccc cctgaccagg gccccggtac gggttcaatt cccatcagtc 120
110 acccgtacac gaaggccccc tccactcgga gggggccttc ggcgttcctg agggttcgcg 180
111 gtcaggcggt cggctcggcg ctggggggact cggccccgtc ggcgggagtg gcctcggcgt 240
112 ccggggga                                                        247
115 <210> SEQ ID NO: 7
116 <211> LENGTH: 255
117 <212> TYPE: DNA
118 <213> ORGANISM: Micromonospora halophytica
120 <400> SEQUENCE: 7
```

RAW SEQUENCE LISTING   DATE: 02/03/2005
PATENT APPLICATION: **US/09/855,340B**   TIME: 09:26:38

Input Set : **A:\PTO.ST.txt**
Output Set: **N:\CRF4\02032005\I855340B.raw**

```
121 tttctccgca cccgcccggg gcgttcgacc gggtgcggcg gcatggtggc tgtagctcag 60
122 ttggcagagc accgggttgt ggtcccggtt gtcgtgggtt caattcccat cagtcacccc 120
123 aggtaagacc caggtcaggg ccggttctca ccggccctga cgcattttca ggggcatggt 180
124 gggggcgcta ccgggggtgg ggtgtctcac cgcgagccag catctcgatc aggcgatcga 240
125 gccggcgctg ccggg                                                255
128 <210> SEQ ID NO: 8
129 <211> LENGTH: 315
130 <212> TYPE: DNA
131 <213> ORGANISM: Micromonospora halophytica
133 <400> SEQUENCE: 8
134 tttctccgca cccgcccggg gcgttcgacc gggtgcggcg gcatggtggc tgtagctcag 60
135 ttggcagagc accgggttgt ggtcccggtt gtcgtgggtt caattcccat cagtcacccg 120
136 gcaagtggat ctactccaca gcagatcagg ccccctccga agagggggcc tgatgcgtca 180
137 taggggacag gtaggggaac tcaacccccg gctccttgct cgcgtcgggt catgccgtcc 240
138 gcgtacccct ccgcgtacct ggccctctcc cgttcctcga tctcggcggc gagctgatcg 300
139 cgcaggtgcg cctcc                                                315
142 <210> SEQ ID NO: 9
143 <211> LENGTH: 260
144 <212> TYPE: DNA
145 <213> ORGANISM: Micromonospora halophytica
147 <400> SEQUENCE: 9
148 taggggaatc cactccggag acgcccggag caatccggag catgacggag caaccagcag 60
149 gtcaggtggc ctgttgaccc cctgaccagg gccccggtac gggttcaatt cccatcagtc 120
150 accccaggta agacccaggt cagggccggt tctcaccggc cctgacgcat tttcaggggc 180
151 atggtggggg cgctaccggg ggtggggtgt ctcaccgcga gccagcatct cgatcaggcg 240
152 atcgagccgg cgctgccggg                                           260
154 <210> SEQ ID NO: 10
156 <211> LENGTH: 209
158 <212> TYPE: DNA
160 <213> ORGANISM: artificial sequence
164 <220> FEATURE:
166 <223> OTHER INFORMATION: pMLP1 attP region
168 <400> SEQUENCE: 10
169 taggggaatc cactccggag acgcccggag caatccggag catgacggag caaccagcag      60
171 gtcaggtggc ctgttgaccc cctgaccagg gccccggtac gggttcaatt cccatcagtc     120
173 acccggcaag tggatctact ccacagcaga tcaggccccc tccgaagagg gggcctgatg     180
175 cgtcataggg gacaggtagg ggaactcaa                                      209
178 <210> SEQ ID NO: 11
180 <211> LENGTH: 19
182 <212> TYPE: DNA
184 <213> ORGANISM: artificial sequence
188 <220> FEATURE:
190 <223> OTHER INFORMATION: primer PR144
192 <400> SEQUENCE: 11
193 tgcttcgacg ccatcargg                                             19
196 <210> SEQ ID NO: 12
198 <211> LENGTH: 20
200 <212> TYPE: DNA
202 <213> ORGANISM: artificial sequence
```

RAW SEQUENCE LISTING   DATE: 02/03/2005
PATENT APPLICATION: US/09/855,340B   TIME: 09:26:38

Input Set : A:\PTO.ST.txt
Output Set: N:\CRF4\02032005\I855340B.raw

```
      206 <220> FEATURE:
      208 <223> OTHER INFORMATION: primer PR145
      210 <220> FEATURE:
      212 <221> NAME/KEY: misc_feature
      214 <222> LOCATION: (7)..(7)
      216 <223> OTHER INFORMATION: n is inosine (I)
      220 <400> SEQUENCE: 12
W--> 221 gtggaanccg ccgaakccgc                                      20
      223 <210> SEQ ID NO: 13
      225 <211> LENGTH: 20
      227 <212> TYPE: DNA
      229 <213> ORGANISM: artificial sequence
      233 <220> FEATURE:
      235 <223> OTHER INFORMATION: primer PDH504
      237 <400> SEQUENCE: 13
      238 agggcaacaa gggaagcgtc                                      20
      241 <210> SEQ ID NO: 14
      243 <211> LENGTH: 21
      245 <212> TYPE: DNA
      247 <213> ORGANISM: artificial sequence
      251 <220> FEATURE:
      253 <223> OTHER INFORMATION: primer PDH505
      255 <400> SEQUENCE: 14
      256 ggcggggggtg tggctattat t                                   21
      259 <210> SEQ ID NO: 15
      261 <211> LENGTH: 21
      263 <212> TYPE: PRT
      265 <213> ORGANISM: artificial sequence
      269 <220> FEATURE:
      271 <223> OTHER INFORMATION: amino acid sequence of open reading frame indicated in
figures 4b
      272         and 4d
      274 <400> SEQUENCE: 15
      276 Ser Pro Asp Ala Glu Ala Thr Pro Ala Asp Gly Ala Glu Ser Pro Ser
      277 1               5                  10                  15
      280 Ala Glu Pro Thr Ala
      281            20
      284 <210> SEQ ID NO: 16
      286 <211> LENGTH: 21
      288 <212> TYPE: PRT
      290 <213> ORGANISM: artificial sequence
      294 <220> FEATURE:
      296 <223> OTHER INFORMATION: amino acid sequence of open reading frame indicated in
figures 5b
      297         and 5d
      299 <400> SEQUENCE: 16
      301 Arg Gln Arg Arg Leu Asp Arg Leu Ile Glu Met Leu Ala Arg Gly Glu
      302 1               5                  10                  15
      305 Thr Pro His Pro Arg
      306            20
```

RAW SEQUENCE LISTING ERROR SUMMARY   DATE: 02/03/2005
PATENT APPLICATION: **US/09/855,340B**   TIME: 09:26:39

Input Set : **A:\PTO.ST.txt**
Output Set: **N:\CRF4\02032005\I855340B.raw**

<u>Please Note:</u>
Use of n and/or Xaa have been detected in the Sequence Listing. Please review the
Sequence Listing to ensure that a corresponding explanation is presented in the <220>
to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:12; N Pos. 7

VERIFICATION SUMMARY   DATE: 02/03/2005
PATENT APPLICATION: **US/09/855,340B**   TIME: 09:26:39

Input Set : **A:\PTO.ST.txt**
Output Set: **N:\CRF4\02032005\I855340B.raw**

L:221 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12 after pos.:0